Self-Organized Spectrum Sharing in Large MIMO Multiple-Access Channels

Romain Couillet
ST-Ericsson, Supélec
630 Route des Lucioles
06560 Sophia Antipolis, France
Email: romain.couillet@supelec.fr

H. Vincent Poor
Department of Electrical Engineering
Princeton University
Princeton, NJ 08544, USA
Email: poor@princeton.edu

Mérouane Debbah
Alcatel-Lucent Chair, Supélec
Plateau de Moulon, 3 rue Joliot-Curie
91192 Gif sur Yvette, France
Email: merouane.debbah@supelec.fr

***Abstract*—In this paper, a deterministic approximation for the rate region of multiple access channels is provided when base station and users have a large number of antennas, and when the transmission bandwidth is divided into several independent subbands. An explicit formulation is also given to the transmit covariance matrices, at each frequency, that reach the boundary of the rate region. From the compact expression of these matrices, suboptimal iterative algorithms emerge that allow the multiple access users to derive autonomously the transmit covariance matrices. This comes at the sole expense of a short signalling overhead, which is constant irrespectively of the number of antennas. Simulations confirm the validity of the theoretical derivations and suggest rather good behaviour obtained by the suboptimal self-organization algorithms.**

I. Introduction

It was proved in [1]-[2] that using multiple antennas in transmit and receive wireless communication devices could lead to significant improvement in terms of achievable transmission rates. Spurred by that promise, the interest in multiple-input multiple-output (MIMO) systems has exploded and many recent technologies for short-to-medium range communications now embed multiple antennas. However, MIMO technologies provide worthwhile gains in performance only at high signal-to-noise ratios (SNRs); such high SNRs are rarely found in outdoor multi-cellular communications, unless the different cells are coordinated and cooperate tightly with each other. Unfortunately, this poses the problem of the large amount of feedback data to be exchanged within cells and between adjacent cells to share the multidimensional frequency-space-user-cell channel information. This is why the idea of self-organization, i.e. local decisions that lead to optimal or close-to-optimal global decisions, has gained interest in fields such as cognitive radios [3].

In this paper, we focus on the particular case of a single-cell multiple access channel (MAC) in which a potentially large number of users, each equipped with many antennas, wish to share a large-band spectrum in a way that maximizes the uplink sum-rate. The general problem is non-trivial and, if solved, would require the base station to feedback to all users the optimal transmit covariance matrices for each subband every time the channel changes. Needless to say, this would be impractical in networks with high mobility. In [4] and [5], Tse and Hanly give an information-theoretic solution of the spectrum sharing problem in the single-antenna case for slow fading and fast fading channels, i.e. they provide an expression of the achievable MAC sum rate as well as an iterative algorithm for the users to reach the optimal transmit (frequency) variance profile. The multi-antenna case is more involved. However, when the number of antennas used at every device is large, it turns out that the sum rate-achieving covariance matrices are asymptotically independent of the local channel variations but depend essentially on the covariance structure of the channel matrix model. This was proven for point-to-point MIMO in Ricean channels, modelled as $\mathbf{A}+\mathbf{X}$ where $\mathbf{A}$ is a deterministic (line-of-sight component) matrix and $\mathbf{X}$ has zero mean independent entries with a variance profile [6], and for multiple access channels when user k-to-base station channels are modelled as Kronecker channels $\mathbf{R}_k^{\frac{1}{2}}\mathbf{X}_k\mathbf{T}_k^{\frac{1}{2}}$, where $\mathbf{X}_k$ is Gaussian, and $\mathbf{R}_k$ and $\mathbf{T}_k$ are deterministic for every k [7].

In this work, we shall describe the rate region of fast fading MACs for wideband MIMO communications when the number of antennas per user and at the base station are large and follow a Kronecker model, and when the transmission bandwidth is divided into multiple independent subbands. We then provide a self-organization method for the users to determine their own optimal transmit signal covariance matrices for every subband, at the sole expense of a scalar feedback to the base station and to the neighboring users, irrespectively of the potentially large number of transmit and receive antennas. When the receive channel correlation matrix at the base station is close to the identity matrix, we will show that the base station does not need to intervene in the self-organization process.

The remainder of this paper is structured as follows: in Section II, we recall the mathematical results from random matrix theory needed here. In Section III, we introduce the MAC MIMO model and derive the optimal users' transmit covariance matrices. In Section IV, the algorithm for self-organization is proposed. In Section V, numerical simulations are carried out. Finally, in Section VI, we draw our conclusions.

Notation: Capital boldface characters denote matrices ($\mathbf{I}_N$ is the $N\times N$ identity matrix). Hermitian transpose is denoted by $(\cdot)^{\mathsf{H}}$. The operator $\det\mathbf{X}$ represents the determinant of matrix $\mathbf{X}$. The symbol $\operatorname{tr}(\mathbf{X})$ denotes the trace of matrix $\mathbf{X}$. The norm

$\|\mathbf{X}\|$ denotes the matrix spectral norm.

II. Preliminaries

We deal in this article with channel models of the type

$$\mathbf{B}_N(\mathcal{K}) = \sum_{k\in\mathcal{K}} \mathbf{R}_k^{\frac{1}{2}}\mathbf{X}_k\mathbf{T}_k\mathbf{X}_k\mathbf{R}_k^{\frac{1}{2}} \tag{1}$$

for a certain set $\mathcal{K}\subset\mathbb{N}$. Each term in (1) originates from a Kronecker channel $\mathbf{R}_k^{\frac{1}{2}}\mathbf{X}_k\mathbf{T}_k^{\frac{1}{2}}\in\mathbb{C}^{N\times n_k}$, i.e. a random matrix $\mathbf{X}_k\in\mathbb{C}^{N\times n_k}$ with independent and identically distributed (i.i.d.) entries with zero mean and variance $1/n_k$, with left and right correlations $\mathbf{R}_k\in\mathbb{C}^{N\times N}$ and $\mathbf{T}_k\in\mathbb{C}^{n_k\times n_k}$, both Hermitian nonnegative definite of bounded normalized trace. As will be shown in Section III, we are interested in the following functional $\mathcal{I}_N(x)$ of $\mathbf{B}_N$:

$$\mathcal{I}_N(x) = \frac{1}{N}\log\det\left(\mathbf{I}_N + \frac{1}{x}\mathbf{B}_N(\mathcal{K})\right). \tag{2}$$

We have shown in [7] that, as N and the elements $\{n_k, k\in\mathcal{K}\}$ grow large, with $N/n_k\to c_k$, such that $0<c_k<\infty$,

$$\mathcal{I}_N(x) - \mathcal{I}_N^\circ(x)\to 0 \tag{3}$$

almost surely, where $\mathcal{I}_N^\circ(x)$ is *deterministic* and defined as

$$\mathcal{I}_N^\circ(x) = \frac{1}{N}\sum_{k\in\mathcal{K}}\log\det\left(\mathbf{I}_N + c_k e_k\mathbf{T}_k\right) + \frac{1}{N}\log\det\left(\mathbf{I}_N + \sum_{k\in\mathcal{K}}\delta_k\mathbf{R}_k\right) - x\sum_{k\in\mathcal{K}}\delta_k e_k \tag{4}$$

with e_i and δ_i satisfying the joint implicit equation

$$\begin{cases} e_i &= \frac{1}{N}\operatorname{tr}\mathbf{R}_i\left(x\left[\mathbf{I}_N + \sum_{k\in\mathcal{K}}\delta_k\mathbf{R}_k\right]\right)^{-1} \\ \delta_i &= \frac{1}{n_i}\operatorname{tr}\mathbf{T}_i\left(x\left[\mathbf{I}_{n_i} + c_i e_i\mathbf{T}_i\right]\right)^{-1}. \end{cases} \tag{5}$$

The value $\mathcal{I}_N^\circ(x)$ is called a *deterministic equivalent* of $\mathcal{I}_N(x)$.[1]

Despite its involved nature, (4) is easily numerically evaluable thanks to a standard fixed-point algorithm to solve (5). When $\mathbf{T}_k$ is replaced by $\mathbf{T}_k\mathbf{P}_k$, where $\mathbf{P}_k$ is an Hermitian nonnegative definite matrix with trace $\operatorname{tr}(\mathbf{P}_k)=P_k$, we also derived in [7] an explicit expression for the $\mathbf{P}_k$, $k\in\mathcal{K}$, which maximizes $\mathcal{I}(x)^\circ$. We will need to generalize this previous result to the case of multiple frequency bands.

III. System Model

A. Transmission model

Consider a wideband multiple access fast fading channel composed of a single base station and of K user terminals. The base station is equipped with N antennas, while user k is equipped with n_k antennas. The communication spectrum is assumed to be shared in F narrow bands sufficiently spaced in frequency, as in cognitive radio resource allocation settings, see e.g. [8]. The uplink channel $\mathbf{H}_{k,f}\in\mathbb{C}^{N\times n_k}$ from user k to the base station in the frequency band $f\in\{1,\ldots,F\}$ is assumed to be flat and is modelled as

$$\mathbf{H}_{k,f} = \mathbf{R}_{k,f}^{\frac{1}{2}}\mathbf{X}_{k,f}\mathbf{T}_{k,f}^{\frac{1}{2}} \tag{6}$$

where $\mathbf{R}_{k,f}^{\frac{1}{2}}$ and $\mathbf{T}_{k,f}^{\frac{1}{2}}$ are the unique nonnegative square roots of the deterministic Hermitian nonnegative matrices $\mathbf{R}_{k,f}$ and $\mathbf{T}_{k,f}$ respectively, and $\mathbf{X}_{k,f}$ is a random matrix with independent and identically distributed (i.i.d.) Gaussian entries of zero mean and variance $1/n_k$. The aforementioned independence of the channels is taken in the sense that the KF matrices $\mathbf{X}_{k,f}$, $1\le k\le K$, $1\le f\le F$, are taken to be independent. In addition, we allow user k to use a maximum total power P_k, distributed over the space-frequency domain; that is, we consider F covariance matrices $\mathbf{P}_{k,1},\ldots,\mathbf{P}_{k,F}$, $\mathbf{P}_{k,f}\in\mathbb{C}^{n_k\times n_k}$ being the transmit signal covariance matrix of user k over frequency band f, satisfying the constraint

$$\sum_{f=1}^{F}\operatorname{tr}\mathbf{P}_{k,f} = P_k. \tag{7}$$

Finally, we assume that the base station is affected by additive white Gaussian noise of variance σ^2 on every antenna.

B. MAC rate region

The fast fading MAC rate region is defined [10] as the union $\mathcal{R}$ of all K-dimensional vectors $(R_1,\ldots,R_k)$, such that, for every subset $\mathcal{K}$ of $\{1,\ldots,K\}$,

$$\sum_{k\in\mathcal{K}} R_k \le \mathrm{E}\left[\frac{1}{N}\sum_{f=1}^{F}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\sum_{k\in\mathcal{K}}\mathbf{H}_{k,f}^{\mathsf{H}}\mathbf{P}_{k,f}\mathbf{H}_{k,f}\right)\right] \tag{8}$$

where the $\mathbf{P}_{k,f}$ satisfy (7) and the expectation is taken over the random $\mathbf{X}_{k,f}$ matrices.

Thanks to the results introduced in Section II, we have that, as the number of antennas N and n_k grow large,

$$\frac{1}{N}\sum_{f=1}^{F}\log\det\left(\mathbf{I}_N + \frac{1}{\sigma^2}\sum_{k\in\mathcal{K}}\mathbf{H}_{k,f}^{\mathsf{H}}\mathbf{P}_{k,f}\mathbf{H}_{k,f}\right) - R^\circ(\mathcal{K})\to 0 \tag{9}$$

almost surely, where we define

$$R^\circ(\mathcal{K}) = \frac{1}{N}\sum_{f=1}^{F}\sum_{k\in\mathcal{K}}\log\det\left(\mathbf{I}_N + c_k e_{k,f}\mathbf{P}_{k,f}\mathbf{T}_{k,f}\right) + \sum_{f=1}^{F}\frac{1}{N}\log\det\left(\mathbf{I}_N + \sum_{k\in\mathcal{K}}\delta_{k,f}\mathbf{R}_{k,f}\right) - \sigma^2\sum_{f=1}^{F}\sum_{k\in\mathcal{K}}\delta_{k,f}e_{k,f} \tag{10}$$

with $e_{k,f}$ and $\delta_{k,f}$ satisfying

$$\begin{cases} e_{k,f} &= \frac{1}{N}\operatorname{tr}\mathbf{R}_{k,f}\left(\sigma^2\left[\mathbf{I}_N + \sum_{k'\in\mathcal{K}}\delta_{k',f}\mathbf{R}_{k',f}\right]\right)^{-1} \\ \delta_{k,f} &= \frac{1}{n_k}\operatorname{tr}\mathbf{T}_{k,f}\left(\sigma^2\left[\mathbf{I}_{n_k} + c_k e_{k,f}\mathbf{P}_{k,f}\mathbf{T}_{k,f}\right]\right)^{-1}. \end{cases} \tag{11}$$

Our first objective is to define the boundary of the polymatroid $\mathcal{R}$, i.e. to find, for every $\mathcal{K}\subset\{1,\ldots,K\}$, those

[1] This can be contrasted with the traditional "asymptotic limit of" found in random matrix theory; the latter does not apply here since, even if all $\mathbf{T}_k$ and $\mathbf{R}_k$ have limiting eigenvalue distributions, $\mathcal{I}_N(x)$ does not necessarily converge.

matrices $\mathbf{P}_{k,l}$, $k \in \mathcal{K}$, which optimize the sum rate (8). More exactly, we will find the $\mathbf{P}_{k,l}$ matrices that maximize the deterministic equivalent (10) of the right-hand side of (8); a reasoning similar to [6] ensures that both *ergodic* capacity and deterministic equivalent coincide asymptotically.

This was pursued in [7] for the particular case $F = 1$. However, we cannot straightforwardly apply these results because (7) introduces a dependence in the power allocation on the different subbands. We need instead to solve the constrained optimization problem

$$\max_{\{\mathbf{P}_{k,f}\}} \mathcal{R}^\circ(\mathcal{K}) \text{ under constraints}$$

$$\forall(k,f),\ \mathbf{P}_{k,f} \geq 0, \text{ and } \forall k,\ \sum_{f=1}^{F} \operatorname{tr}(\mathbf{P}_{k,f}) = P_k. \quad (12)$$

For this, first consider (10) as the image of the following function

$$V : (\{\mathbf{P}_{k,f}\}_{k,f}, \{e_{k,f}\}_{k,f}, \{\delta_{k,f}\}_{k,f}) \mapsto R^\circ(\mathcal{K}) \quad (13)$$

in the $3F|\mathcal{K}|$ variables $e_{k,f}$, $\delta_{k,f}$ and $\mathbf{P}_{k,f}$. We then have

$$\frac{\partial R^\circ(\mathcal{K})}{\partial \mathbf{P}_{k,f}} = \sum_{k',f'} \left(\frac{\partial V}{\partial e_{k',f'}} \frac{\partial e_{k',f'}}{\partial \mathbf{P}_{k,f}} + \frac{\partial V}{\partial \delta_{k',f'}} \frac{\partial \delta_{k',f'}}{\partial \mathbf{P}_{k,f}} \right) + \frac{\partial V}{\partial \mathbf{P}_{k,f}} \quad (14)$$

where

$$\frac{\partial V}{\partial e_{k,f}} = \sum_{f=1}^{F} \left[\frac{1}{N} \operatorname{tr} \mathbf{T}_{k,f} \left(\mathbf{I}_{n_k} + c_k e_{k,f} \mathbf{P}_{k,f} \mathbf{T}_{k,f} \right)^{-1} - \sigma^2 \delta_{k,f} \right] \quad (15)$$

and

$$\frac{\partial V}{\partial \delta_{k,f}} = \sum_{f=1}^{F} \left[\frac{1}{n_k} \operatorname{tr} \mathbf{R}_{k,f} (\mathbf{I}_N + \sum_{k' \in \mathcal{K}} \delta_{k',f} \mathbf{R}_{k',f})^{-1} - \sigma^2 e_{k,f} \right]. \quad (16)$$

Going back to our initial problem, $e_{k,f}$ and $\delta_{k,f}$ are defined by the implicit equation (11). This implies that the above derivatives equal zero. As a consequence,

$$\frac{\partial R^\circ(\mathcal{K})}{\partial \mathbf{P}_{k,f}} = \frac{\partial V}{\partial \mathbf{P}_{k,f}}. \quad (17)$$

The dependence on $\mathbf{P}_{k,f}$ in V is reduced to the term $\frac{1}{N} \sum_{f=1}^{F} \log\det(\mathbf{I}_N + c_k e_{k,f} \mathbf{P}_{k,f} \mathbf{T}_{k,f})$. Noticing that V, as a function of the joint variable $\{P_{k,f}\}_{k,f}$ is concave, optimizing along $\{P_{k,f}\}_{k,f}$ is equivalent to optimizing independently along each $\mathbf{P}_{k,f}$. The optimal $\mathbf{P}_{k,f}$ for all k, f are then solutions of a standard water-filling algorithm. For ease of reading, we add the superscript '$\star$' to the parameters linked to the solution that maximizes the deterministic equivalent for the ergodic capacity; e.g. $\mathbf{P}^\star_{k,f}$ denotes the transmit covariance $\mathbf{P}_{k,f}$ which maximizes (10). We then have $\mathbf{P}^\star_{k,f} = \mathbf{U}_{k,f} \mathbf{Q}^\star_{k,f} \mathbf{U}^{\mathsf{H}}_{k,f}$, where $\mathbf{U}_{k,f} \in \mathbb{C}^{n_k \times n_k}$ is a matrix whose columns are the n_k eigenvectors of $\mathbf{T}_{k,f}$ and $\mathbf{Q}^\star_{k,f}$ is diagonal with i^{th} diagonal entry $q^\star_{k,f,i}$ satisfying

$$q^\star_{k,f,i} = \left(\mu_k - \frac{1}{c_k e^\star_{k,f} t_{k,f,i}} \right)^+ \quad (18)$$

where $t_{k,f,i}$ is the i^{th} eigenvector of $\mathbf{T}_{k,f}$ and $\mu_k > 0$ is such that $\sum_{f,i} q^\star_{k,f,i} = P_k$.

The main consequence of (18) is the surprising fact that, for sufficiently large N, n_k, the optimal power allocation strategy depends on the transmit correlation eigenvalues $t_{k,f,i}$ weighted by a *unique* scalar $e^\star_{k,f}$ for every frequency f. This means that, if user k is embedded with a large number of antennas, then user k can determine on its own the transmit matrices $\mathbf{P}^\star_{k,1}, \ldots, \mathbf{P}^\star_{k,F}$ if it is provided with the F scalars $e^\star_{k,1}, \ldots, e^\star_{k,F}$. The next section is dedicated to determining algorithms for every user to obtain these $e^\star_{k,f}$'s parameters.

IV. Self-Organized MIMO MAC

A. Base station-supported iterative water-filling

We shall present in the following a method for user k to autonomously identify the F optimal $n_k \times n_k$ signal covariance matrices $\mathbf{P}^\star_{k,1}, \ldots, \mathbf{P}^\star_{k,F}$. Obviously, if the base station is aware of all $\mathbf{R}_{k,f}$ and $\mathbf{T}_{k,f}$, then it can process the water-filling formula (18) and broadcast the FK optimal power matrices to the users; this however demands $F \sum_{k=1}^{K} n_k^2$ scalar transmissions, which may form a large signalling overhead, if K and the n_k's are large. As previously mentioned though, from the viewpoint of user k, the entries $q^\star_{k,f,i}$ for all (f,i) in (18) depend on the $t_{k,f,i}$, for all (f,i) which user k might be aware of, and of the $e^\star_{k,f}$, for all f, which user k *cannot* compute unless it knows all the $\mathbf{R}_{k',f}$, $\mathbf{T}_{k',f}$ and $\mathbf{P}_{k',f}$ for all (k',f) from the other users. We wish therefore to design a cooperative process between the base station and the users for the latter to obtain their respective values $e^\star_{k,f}$; this process must be minimal in terms of peer-to-peer signalling. We will in fact show that only $O(FK)$ scalar values need to be successively exchanged throughout the network for a good approximation of the optimal covariance matrices to be recovered by every user.

First, we recall from [7] that, if all $\mathbf{R}_{k,f}$ and $\mathbf{T}_{k,f}$ are known (say, at the base station), then an *iterative water-filling* algorithm can be implemented which, upon convergence, converges surely to the optimal $\mathbf{P}_{k,f}$ matrices.

The iterative algorithm consists in (i) initializing the values for $e_{k,f}$ and $\delta_{k,f}$ to given constants, and (ii) successively computing (until convergence),

- for fixed $e_{k,f}$ and $\delta_{k,f}$, the entries $q_{k,f,i}$ (and therefore the $\mathbf{P}_{k,f}$) that solve (18), and
- for fixed $q_{k,f,i}$, the updated $e_{k,f}$ and $\delta_{k,f}$ that satisfy (11).

Based on the above iterative water-filling algorithm, we propose a self-organizing scheme for which we make the reasonable assumptions that the base station knows at least the $\mathbf{R}_{k,f}$ matrices, for all (k,f), and that user k knows at least its own matrices $\mathbf{T}_{k,f}$, for all f. Our scheme basically lets the base station first update the $e_{1,f}$, for all f, and transmit the updated values to user 1. Upon reception, user 1 updates the $q_{1,f,i}$ for all (f,i), for all f,i, and updates the $\delta_{1,f}$ for all f. These are transmitted back to the base station which now computes the updated $e_{2,f}$ for all f, transmitted back to user 2; the process continues up to user K, and then back to user 1 until convergence is reached. Formally, the algorithm unfolds as in Algorithm 1.

Algorithm 1 Base station-aided iterative water-filling

Initialization: for all k, f, $\delta_{k,f} = 1$. Define convergence threshold $\varepsilon > 0$.
while $\max_{k,f} \|\mathbf{P}_{k,f} - \mathbf{P}^\star_{k,f}\| > \varepsilon$ **do**
  for $k \in \{1, \ldots, K\}$ **do**
    for $f \in \{1, \ldots, F\}$ **do**
      The base station computes $e_{k,f}$ from (11)
    end for
    The base station transmits $(e_{k,1}, \ldots, e_{k,F})$ to user k
    for $f \in \{1, \ldots, F\}$ **do**
      Based on $e_{k,f}$, user k computes $\mathbf{P}_{k,f}$ from (18)
      Based on $e_{k,f}$ and $\mathbf{P}_{k,f}$, user k computes $\delta_{k,f}$
    end for
    User k transmits $(\delta_{k,1}, \ldots, \delta_{k,F})$ to the base station
  end for
end while

The steps of Algorithm 1 do not follow those of the iterative water-filling process proposed in [7] when the knowledge on the $\mathbf{R}_{k,f}$ and $\mathbf{T}_{k,f}$ for all f is centralized at the base station; the convergence *to the correct solution* cannot as a consequence be ensured, unless the $e_{k,f}$ and $\delta_{k,f}$ are passed along on several loops before updating the $\mathbf{P}_{k,f}$. It is indeed shown in [6] that, under this condition, if the algorithm converges, then it converges to the correct solution. This requires more processing delay though and, as simulations confirm, barely leads to significant improvement unless the transmit correlation pattern is strong.

B. Self-organized iterative water-filling

Now, observe that, when there is no correlation at the base station, i.e. for all (k, f), $\mathbf{R}_{k,f} = \mathbf{I}_N$, then the K users can autonomously determine the $\mathbf{P}^\star_{k,f}$, without the need for feedbacks to and from the base station. Indeed, (11) becomes

$$\begin{cases} e_f^{-1} &= \sigma^2 \left(1 + \sum_{k'=1}^K \delta_{k',f}\right) \\ \delta_{k,f} &= \frac{1}{n_k} \operatorname{tr} \mathbf{T}_{k,f} \left(\sigma^2 \left[\mathbf{I}_{n_k} + c_k e_f \mathbf{P}_{k,f} \mathbf{T}_{k,f}\right]\right)^{-1} \end{cases} \tag{19}$$

where e_f replaces $e_{k,f}$, which is now independent of k.

Assume all e_f are initialized to constants known to user 1. Now the situation is the following: user 1 computes all $\mathbf{P}_{1,f}$, then computes the resulting $\delta_{1,f}$ and finally updates the e_f. Those e_f are transmitted *directly* to user 2. User 2 now computes the $\mathbf{P}_{2,f}$ based on the e_f received by user 1, call them $e_f^{(\mathrm{old})}$, then updates the $\delta_{2,f}$ stored in memory, call them $\delta_{2,f}^{(\mathrm{old})}$, into new $\delta_{2,f}$. The e_f are finally updated as

$$e_f^{-1} = e_f^{(\mathrm{old})} + \sigma^2 (\delta_{2,f} - \delta_{2,f}^{(\mathrm{old})}) \tag{20}$$

This new e_f is then transmitted to user 3 and the algorithm goes on until convergence. This can be written in the form of Algorithm 2.

Once again, if Algorithm 2 converges, convergence to the correct solution is not true in general but simulations will suggest a close-to-optimal behaviour. Asynchronous signalling techniques are currently being studied for faster convergence of Algorithms 1 and 2.

Algorithm 2 Self-organized iterative water-filling

Initialization: for all k, f, $\delta_{k,f} = 1$. Define convergence threshold $\varepsilon > 0$.
while $\max_{k,f} \|\mathbf{P}_{k,f} - \mathbf{P}^\star_{k,f}\| > \varepsilon$, **do**
  for $k \in \{1, \ldots, K\}$ **do**
    for $f \in \{1, \ldots, F\}$ **do**
      Based on e_f, user k computes $\mathbf{P}_{k,f}$ from (18)
      Based on $\{e_f, \mathbf{P}_{k,f}\}$, user k computes $\delta_{k,f}$ from (19)
      Based on $\delta_{k,f}$, user k updates $e_{k,f}$ from (19)
    end for
    User k transmits $(e_{k,1}, \ldots, e_{k,F})$ to user $k + 1 \pmod K$
  end for
end while

Notice that this message passing approach may require several loops over all users to converge, in which case the signaling overhead incurred is not beneficial compared to direct feedback from the base station. However, from a position of equilibrium, users that experience new channel correlation patterns can initiate on their own the transmission of their updated e_f to neighboring users. The updating process is then necessarily fast converging, timely, and is consistent with the current cognitive radio incentive, which aims at transferring part of the network intelligence burden to terminal users.

V. Simulations and Results

In this section, we present numerical simulations of the theoretical results and algorithms presented before. Our case study is that of a two-user MAC, each user equipped with $n_1 = n_2 = 8$ antennas while the base station has $N = 8$ antennas. The number F of subbands is $F = 2$, and the SNR is 20 dB. The matrices $\mathbf{T}_{k,f}$ and $\mathbf{R}_{k,f}$ originate from both the contribution of (i) a particular antenna spacing of the linear antenna array model, and (ii) a particular solid angle of energy arrival/departure. From these features, we consider Jake's model to compute the entries of the covariance matrices. For instance, the $(a, b)^{th}$ entry of any $\mathbf{T}_{k,f}$ has the form

$$\int_{\theta_{\min}}^{\theta_{\max}} \exp\left(2\pi i \frac{d_{ab}}{\lambda} \cos(\theta)\right) d\theta \tag{21}$$

where d_{ab} denote the distance between antennas a and b, i.e. $|b - a|$ times the wavelength λ; and where $\theta_{\min}$ and $\theta_{\max}$ are the minimum and maximum energy spreading angles, whose absolute difference is $\pi/2$.

In Figure 1, we first compute the optimal $\mathbf{P}^\star_{k,f}$ and compare (i) the effective rate region, evaluated from a similar descent algorithm as in [9] and averaged over $1,000$ channel realizations, (ii) the theoretical deterministic equivalent (10), (iii) the region obtained from Algorithm 1 with stopping time threshold $\varepsilon = 10^{-2}$, and (iv) the rate region for uniform power

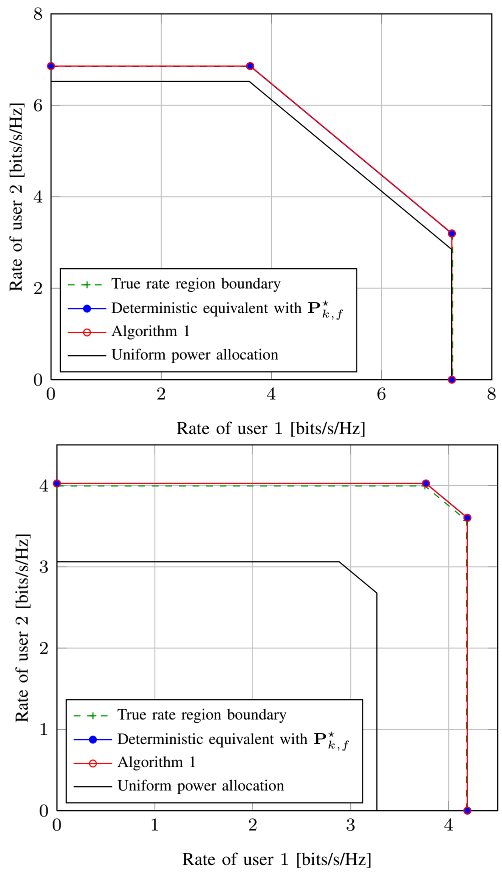

Fig. 1. Rate region of a two-user MAC, $N = 8$, $n_1 = n_2 = 8$, $SNR = 20$ dB, $F = 2$, $\mathbf{T}_{k,f}$ and $\mathbf{R}_{k,f}$ determined by inter-antenna distance and solid angle of energy departure/arrival, $P_1 = P_2 = 1$. Top: case (a), low transmit antenna correlation, bottom: case (b), high transmit antenna correlation.

allocation across transmit antennas. Denote λ the transmit signal wavelength. We consider two cases: (a) low correlation, for which the distance between adjacent transmit antennas is 10λ at the base station and λ at both users, and (b) high correlation, for which the distance between transmit antennas is λ at the base station and 0.1λ at both users. The solid angles of signal arrival/departure are limited to $\pi/2$ in the horizontal plane at both the receiver and the transmitters for varying $\theta_{\min}$ and $\theta_{\max}$. We observe that, even for these small values of N, n_k, the approximated rate region is extremely close to the true averaged rate region. As for the *suboptimal* Algorithm 1, it shows a surprisingly accurate performance, either in low or high correlation.

Note additionally that, in both for cases (a) and (b), the number of iterations required for Algorithm 1 to converge were of order 4 to 5 iterations only. The total number of signalling feedback in these cases amounts to approximately 20 real scalar values to be exchanged. If the base station were to broadcast the full covariance matrices to the users, it would require 64 real values per user per subband to be transmitted, so a total of 256 real values to be exchanged. This eventually corresponds to a ten-fold division of the feedback overhead, when all users start from a uniform power allocation policy. Further simulations for $N = n_1 = n_2 = 32$ show that Algorithm 1 still converges in a maximum of 5 iterations; in this case the corresponding feedback gain is 20 reals against 4096 reals for full matrix transmission, hence a 200-fold gain.

Note finally that those simulations assume rather random initial values for $e_{k,f}$, $\delta_{k,f}$ etc. When Algorithm 1 is employed on field, it must be run any time some user enters or leaves the network or when the long-term correlation coherence time has expired. In any of these situations, initializing Algorithm 1 to the previously optimal $e_{k,f}$, $\delta_{k,f}$ etc. should reduce the number of iterations required for convergence, making the feedback gain even more dramatic.

VI. Conclusion

We have provided a deterministic equivalent to the rate region of correlated MIMO wideband MACs, when the wideband channel is subdivided into several independent subbands. We have determined the transmit space-frequency covariance matrices that maximize the uplink sum rate, under a per-user power constraint. Additionally, we have provided two suboptimal algorithms for the users to determine their optimum power allocations in a decentralized manner, either (i) by exchanging feedback signals with the base station or, (ii) under limited antenna correlation at the base station, by exchanging signals with the neighboring users. Both strategies have been shown to generate very little transmission overhead compared to the centralized approach, and have also been shown by simulations to converge rapidly to a solution which, although theoretically suboptimal, is indistinguishable from the optimal power allocation.

References

[1] G. J. Foschini and M. J. Gans, “On limits of wireless communications in a fading environment when using multiple antennas,” *Wireless Personal Communications*, vol. 6, no. 3, pp. 311-335, 1998.

[2] I. E. Telatar, “Capacity of multi-antenna Gaussian channels,” *European Transactions on Telecommunications*, vol. 10, no. 6 pp. 585-595, 1999.

[3] J. Mitola III, and G. Q. Maguire Jr, “Cognitive radio: Making software radios more personal,” *IEEE Personal Communications*, vol. 6, no. 4, pp. 13-18, 1999.

[4] D. N. C. Tse, and S. V. Hanly, “Multiaccess fading channels. I. Polymatroid structure, optimal resource allocation and throughput capacities,” *IEEE Trans. Information Theory*, vol. 44, no. 7, pp. 2796-2815, 1998.

[5] S. V. Hanly, and D. N. C. Tse, “Multiaccess fading channels. II. Delay-limited capacities,” *IEEE Trans. Information Theory*, vol. 44, no. 7, pp. 2816-2831, 1998.

[6] W. Hachem, Ph. Loubaton, and J. Najim, “A CLT for information theoretic statistics of Gram random matrices with a given variance profile,” *Annals of Applied Probability*, vol. 18, no. 6, pp. 2071-2130, 2008.

[7] R. Couillet, M. Debbah, and J. W. Silverstein, “A Deterministic Equivalent for the Capacity Analysis of Correlated Multi-User MIMO Channels,” *submitted to IEEE Trans. Information Theory*, ArXiv Preprint http://arxiv.org/abs/0906.3667

[8] A. Leshem, and E. Zehavi, “Game theory and the frequency selective interference channel - A tutorial,” Arxiv preprint arXiv:0903.2174, 2009.

[9] M. Vu, A. Paulraj, “Optimal linear precoder for MIMO wireless correlated channels with nonzero mean in space-time coded systems,” IEEE Trans. on Signal Processing, vol. 54, no. 6, pp. 2318, 2006.

[10] A. Paulraj, R. Nabar and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge University Press, Cambridge, ULC, 2003.